DOCUMENT RESUME

ED 270 500 TM 860 390

AUTHOR Burt, Charles W.; Geis, George L.
TITLE Guidelines for Developmental Testing: Proposed and Practiced.
PUB DATE Apr 86
NOTE 75p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Data Collection; *Educational Practices; Elementary Secondary Education; *Evaluation Methods; Evaluation Utilization; Experimental Groups; Foreign Countries; *Formative Evaluation; Higher Education; *Instructional Design; *Instructional Development; *Instructional Material Evaluation; Material Development; Questionnaires; Sampling; Surveys
IDENTIFIERS Canada; *Developmental Testing

ABSTRACT

A survey was conducted of expert instructional design practitioners about their use of developmental testing and revision procedures. The degree of correspondence was examined between published recommended practice and actual practice, as reported by 90 respondents to a questionnaire. The surveyed sample were members of the National Society for Performance and Instruction. The SPSSX Multiple Response procedure was used to analyze the questionnaire data, since respondents were permitted to select more than one response per question. In developmental testing, which is a form of formative evaluation, qualitative or quantitative data, or both, are obtained for the revision of the instructional product. Results indicated that respondents endorsed a number of the recommended practices, such as using experimental subjects, using peer review, using both criterion-referenced tests and subjects' comments, and using subjects of varying abilities. Based on the response patterns found in the data analysis, the following representative profile of developmental testing and revision procedures is presented: (1) select a representative sample; (2) conduct small group sessions and perhaps also one-to-one sessions; (3) pretest the subjects, present the learning material, note their responses, posttest, and debrief the subjects; (4) analyze and interpret the data; (5) make revisions; and (6) test the revisions. (The questionnaire is appended). (GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED270500

Guidelines for Developmental Testing: Proposed and Practiced

Charles W. Burt & George L. Geis

McGill University

A round table paper presented at the annual meeting
of the American Educational Research Association,
San Francisco, April 1986

Preparation of this paper was supported by the Quebec Programme de formation de chercheurs et aide a la recherche (FCAR), Grant #290-10.

The authors gratefully acknowledge the contributions of the members of the research team: Cynthia B. Weston, Tony Mark and Alenoush Saroyan.

TM 860 370

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C. W. Burt

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

There is strong consensus, on the part of proponents of systematic instructional design models, that components within a model should reflect an iterative cycle of evaluation for revision purposes. Many refer to this process as formative evaluation since it is conducted during the development of a product for the purpose of improving the product through revision. Andrews & Goodson (1980) found that the formative evaluation process was included in 38 out of the 40 instructional design models they reviewed. Figure 1 presents an instructional design model not unlike many found within the literature. Instead of referring to the activity as formative evaluation, this model uses the terms developmental testing and revision - recognizing the important historical precedent set by the literature on programmed instruction which used this term (e.g., Horn 1966). The authors of this paper consider the term formative evaluation to be too broad; while it includes the concept of developmental testing it is also commonly used in the literature to subsume field testing and program evaluation.

Several issues and future research questions will be discussed within this paper which an instructional design research team at McGill University has been active in addressing. The results of the questionnaire reported here are only part of a systematic research programme. The first part of the programme, the literature review and isolation of developmental testing and revision variables, was reported at AERA in 1984, (Geis, Burt, & Weston, 1984). The second part of the programme, results of a survey of "expert" instructional design practitioners' endorsement of developmental testing and revision procedures, is embodied within the presentation of this paper. A third part of the research programme, experimentally investigating qualitative and quantitative evaluation methodologies under various conditions, is presently underway.

This paper is concerned with variables related to developmental testing and revision. It is not the purpose of this presentation to review research on these variables. Rather, the paper examines the degree of correspondence between recommended practice (taken from 17 leading instructional design texts) and actual practice (the results of a large survey of "expert" instructional designers) during the formative evaluation process.

The identification and classification of developmental testing and revision variables will be described as well as how a questionnaire used to survey practitioners was constructed, pilot tested, and distributed. The kinds of analyses used to treat the data and the results of the analyses will be reported and discussed. A general profile of "expert" practices will be presented. Directions for future research are suggested.

Analysis and Classification of Variables

As a result of reviewing the textbook literature on systematic instructional design and the related literature on formative evaluation (focusing on the developmental testing and revision of prototype instructional products), an annotated bibliography of recommendations was compiled. This bibliography represents 16 textbooks and one dissertation that devote substantial discussion to developmental testing and revision activities. In essence, the literature is composed of a large number of procedures that are recommended to practitioners of instructional design. Table 1 reflects a synthesis of this textbook literature.

A content analysis of the various procedures (rule statements) provided the basis for the construction of a classification system for all of the procedures being endorsed by authors within the sampled literature. This permitted the "mapping" of the range of recommendations between authors and highlighted the conflicting nature of the procedures. All of the authors, in differing ways, focused their attention on the following variables:

1. The criteria for acceptable student performance.
2. The number of test subjects (learners) required during developmental testing.
3. The characteristics of the test subject population.
4. The role of the test subject(s) during developmental testing and revision.
5. The role of the instructional designer/evaluator during developmental testing and revision.
6. The method of collecting data.
7. The basis for revision.
8. The testing environment.

Questionnaire Construction

All of the rule statements gleaned from the literature were classified in order to facilitate cross-comparisons. Essentially, this permitted the isolation of the critical variables involved in the developmental testing and revision process. The results of this review and analysis of the literature was reported in two paper presentations at AERA by Geis, Burt, & Weston (1984) and Weston, Burt, & Geis (1984). The isolation of the critical variables facilitated the construction of the first draft of a questionnaire designed to solicit respondents' agreement with the various rule statements found in the literature. In addition, respondents were asked to provide a justification for their chosen procedures. This was included because the literature review revealed a paucity of justifications for many of the procedures authors were advocating.

Eleven "experts" in the area of instructional design were sent a first draft of the questionnaire and asked to respond to it while noting comments, questions, and suggestions on format and content. A follow-up telephone interview with each respondent (unstructured) was conducted for about 10 minutes approximately 10 days after its mailing. All of their comments were compiled and common trends in the suggested revisions were identified. As a result, the following revisions were made:

- A cover letter for the final draft of the questionnaire added a general definition of both developmental testing and revision. Respondents were directed to an additional section in the survey if they could not accept the general definitions; this allowed respondents to explain their unique procedures. Another section was added to the questionnaire to enable respondents to describe how they would test and revise if not operating under practical constraints, e.g., time and budget. With this question providing "ideal" procedures, it was thought that it would be useful to compare respondents' typical procedures with what they felt would be ideal.
- Sections soliciting demographic information were deleted because it was found that readily accessible demographics on the sample were available from the National Society for Performance and Instruction's Membership Directory.
- Another response set was added to permit respondents to sequence their responses (activities) on a time line.

- a question attempting to identify whether or not designers have found that successive test subjects generally replicate each others' comments was added.
- Some of the questions originally included responses to both testing and revision. A question having this problem was divided in two, one for testing and one for revision.

The final version of the questionnaire and its accompanying cover letter (see Appendix 1) focused on asking respondents (practitioners in the field) to select the recommended practices to which they in fact adhered. The final version of the questionnaire required respondents to:

- characterize the test subject population.
- explain how test subjects were differentiated with regards to entry skills, I.Q., etc.
- indicate how many test subjects were typically used.
- elaborate on the role of the test subject and the designer during developmental testing and revision.
- report when testing and revision was terminated.
- explain how data was obtained during developmental testing and which procedures provided the most useful data for revision.
- indicate how collected data was analyzed, to provide a basis upon which instruction was revised.
- indicate how explicit and teachable the respondents' procedures were for testing and revision. Respondents were also asked to provide documentation, if existing, for these procedures.

For all questions posed, respondents could utilize an "OTHER" response to identify procedures that the questionnaire did not address. After selecting/identifying their procedures, respondents were asked to provide a rationale, in open-ended format, for their utilization of the specified procedures. All respondents were also requested to identify, in open-ended format, their "ideal" procedures, i.e., as if there were no timeline or budget constraints operating.

Procedures for data collection

It was decided that the sample of experts to receive the questionnaire should be experts in the area of systematic instructional design. The National Society for Performance and Instruction was chosen as the source from which to build a pool of possible respondents. This international professional organization is well known for its endorsement of systematic instructional design principles for the production of instruction. It was decided to survey all of NSPI's presenters from their 1984 annual meeting, held in Atlanta, Georgia. A Participant Directory of the meeting included the names and addresses, job descriptions, research interests, and other demographics of all of the presenters at the conference. In addition, it suggested that a great many of the presenters would be able to provide valid responses to the questionnaire because of their expressed interest in using systematic instructional design procedures. On this basis, each presenter was sent a questionnaire. If a presenter had already participated in the initial pilot testing of the questionnaire or if a presenter's textbook had been taken into account for the compilation of the questionnaire, they were dropped from the sample.

On July 5, 1984, 206 participants were sent a questionnaire. Mailing procedures did not attempt to provide anonymity for the sample. On September 11, 1984, 115 non-respondents were sent out a reminder card. Table 2 reports the final response rate for the various categories of the sample. Some of the sample returned an uncompleted question because they were not involved in instructional design. Some reported that they used a completely different approach that did not permit them to fill out the questionnaire. Part of the sample did not return the questionnaire. Table 2 has been constructed so that those who in some way responded can be compared to the total sample (N=206). There was a total response rate to the questionnaire of 62%. Table 3 presents a demographic breakdown of the samples' principal type of employment both within and across sample response-type categories, i.e., total sample; respondents; non-respondents; and those not involved in instructional design. Table 4 presents a similar demographic breakdown for the samples' most significant employment responsibility. For both of these important demographic variables, there were no practical differences across sample

response-type categories. This does suggest that the respondents used in the data analysis (N=90) were probably representative of the total sample (N=206).

Data analysis

Two types of data emerged from the questionnaire. For open-ended response sets, i.e., the rationales for selecting procedures and the "ideal" procedures, a content analysis was performed. This paper does not present these results but they will be reported at a later date. The paper does present the results of the objective data coded from the questionnaire and analyzed using the SPSSX Multiple Response procedure - as respondents (N=90) were permitted to select more than one response per question. Nonparametric tests of significance (Chi-square), when appropriate, were used to reveal significant differences between proportions of responses in order to test various research hypotheses.

Discussion of results

The purpose of developmental testing is to obtain data about what needs to be revised in an instructional product. Various kinds of information can be generated, e.g., criterion test data, test subject comments, and observations of non-verbal behavior. Essentially two different kinds of data can be collected: qualitative and/or quantitative. In all cases, the purpose is to obtain data for the revision of the instructional product. Some authors in the literature emphasize reliance on quantitative test data while others suggest reliance on both quantitative and qualitative feedback. Decisions about the kinds of data to be obtained will obviously influence the ways in which the sessions are conducted, i.e., the roles of the designer/evaluator and the test subjects. The purpose of the data collection activities will complement the roles of both the evaluator and test subject.

The results of the data analysis will be reported in terms of a comparison between the recommendations from the literature and the reported practices of the sampled practitioners. The recommendation from the literature will be presented first and the reported practice (questionnaire finding) second. The two will be compared.

Recommendation: It is virtually universally suggested to use test subjects (learners) in the developmental testing of proto-type instructional materials.

Finding: All respondents to the questionnaire indicated that they used test subjects. Eight respondents did indicate that instead of using test subjects, subject matter experts or peer reviews served instead. These respondents then went on to qualify this particular response by stating that these other review mechanisms were utilized only when test subjects were unavailable or when timelines pre-empted the use of test subjects. This finding strongly concurs with the literature's recommendation to use test subjects.

Recommendation: The literature consistently endorses the use of peer reviews in addition to using test subjects.

Finding: A significant proportion of respondents endorsed this procedure as well.

Recommendation: Some authors recommend revising during the session; some after the session; and others suggest revising at any time. There is no strong concensus.

Finding: A significant proportion of the practitioners indicated that as a general strategy, revision takes place after the testing session.

Recommendation: Most authors indicate that revisions should be based on both criterion test data and test subjects' comments.

Finding: A significant proportion of the respondents concurred with the recommendation from the literature. The samples' responses, taken as a whole, does indicate that some practitioners place more emphasis on criterion test results than test subject comments and vice versa. Three kinds of approaches to evaluation emerge: quantitative, qualitative, and both quantitative and qualitative. Those practitioners who rely solely on criterion test data (15.9% of respondents)

take a quantitative position; those who focus solely on test subjects' comments (25.6% of respondents) take a qualitative approach; and those who utilize both criterion test data and students' comments (58.5% of respondents) are eclectic in their procedures.

Recommendation: The literature strongly suggests collecting attitudinal data, usually by means of a formal attitudinal survey instrument.

Finding: A significant proportion of the respondents do not administer an attitudinal survey.

Recommendation: Most authors advocate that the test subjects used during developmental testing should be characterized as being representative of the target population for whom the instruction was designed. This usually means a mix of high, medium, and low level/aptitude subjects.

Finding: A significant proportion of the survey sample agreed that test subjects, in some way, should represent high, medium, and low level/aptitude performers.

Recommendation: Related to the above recommendation, authors diverge on suggestions for differentiating test subject level/ability. Some recommend pretest scores, others standardized aptitude tests, and still others past achievement data, etc. The literature is by no means clear as to what the characteristics of test subjects should optimally be nor how to select individuals.

Finding: A significant proportion of the respondents indicated that in order to differentiate test subject level/aptitude before the session, pre-existing data (e.g., a standardized test, past achievement data, job performance measures, etc.) was utilized as opposed to having the instructional designer/evaluator generate this kind of information (e.g., pretest results).

Recommendation: There is extreme variability when authors suggest the number of subjects to be utilized during developmental testing. Some recommend conducting clinical one-to-one sessions only; some recommend small group sessions; and some recommend both one-to-one and

small group. Authors vary as to the number of one-to-one sessions to be conducted; the number of small groups; as well as to the number of test subjects within a small group. Generally the literature asserts that several test subjects be used and/or several types of sessions. This implies that there is some variability in test subject generated data across sessions.

Finding: The surveyed practitioners' responses indicate that in practice there is no significant difference in the proportions using solely small group sessions as opposed to both one-to-one and small group sessions. Only eight respondents indicated that they used just one-to-one data. When directly asked to provide the actual number of sessions and test subjects for both one-to-one and small group conditions, the practitioners tended to respond by providing ranges, e.g., 2 - 3 test subjects. The mean responses using the minimum and maximum values of the ranges are:

a) number of one-to-one sessions: min. = 3.2/ max. = 4.6

b) number of small group sessions: min. = 2.1 / max. = 3.4

c) number of test subjects typically used in each small group session:

min. = 7.1 / max. = 10.8

Respondents were asked if test subjects, during successive developmental testing and revision sessions, a) generally provided different kinds of information, b) different subjects generally replicated each other, or c) both of these conditions had been observed depending on the context. A significant proportion indicated that different test subjects often provide qualitatively different information during successive sessions. This could indicate a rationale for using several test subjects and more than one type of session. By varying the number of test subjects and sessions, much of the variability of test subject generated data supposedly can be identified and addressed.

Recommendation: Keeping in mind the qualitative/quantitative distinction with regard to the kind of data that can be collected, the literature offers numerous suggestions on ways to obtain the data during developmental testing. These include:

- scores from criterion pretest and posttest
- responses to an attitudinal survey

- responses to the probing (during the test session) of test subjects
- recorded test subject behavior (by note-taking, audio tape, or video tape)
- data from debriefing test subjects (after the test session)
- test subject notations written on the instructional materials
- test subject comments placed on a revision comment sheet
- notes that were taken on examples, illustrations, information added, and changes to sequences made while the test subject was tutored through the instructional material
- spontaneous verbal comments that the test subjects have made during the session

Although by no means consistent in making recommendations (as can be seen from the above suggestions), instructional design textbook authors generally recommend a mix of both qualitative and quantitative data collection strategies. It appears that most would recommend collecting pre- and posttest data, attitudinal data, test subject generated comments while going through the material, designer/evaluator observational data, and debriefing (interview) data.

Finding: When asked directly: "How do you obtain your data during developmental testing?", respondents most frequently indicated:

- from criterion pre- and posttests
- by recording information (by note-taking, audio tape, or video tape)
- by debriefing test subjects after the session
- from test subject notations written on the instructional materials
- from spontaneous verbal comments that the test subjects have made during the test session

When asked to indicate the <u>most useful</u> of all of the varying procedures, respondents chose three strategies. These three strategies were then rank ordered by their frequency of perceived utility as revision data sources. In rank order, they are:

1. By debriefing test subjects (after the session)
2. From criterion pretest and posttests
3. From test subject notations written on the instructional materials

Recommendation: As previously mentioned, the kinds of data that an evaluator decides to collect will affect the roles that both the evaluator and test subject will play during both developmental testing and revision. In general, a majority of the authors suggest that both evaluator and test

subject be encouraged to play an active role during the testing session. Further interaction is facilitated by debriefing after the session - the data can be discussed in an interview situation. The results of the testing determine the agenda.

Finding: A large majority (81/90) of the respondents place the test subject into an active role during developmental testing. This role permits the test subjects to freely comment on problems with the instructional materials. Raw frequencies reveal that the role of a test subject during testing is to ask questions or make comments. It is the designer/evaluator's role to encourage the test subject to do so. The majority of respondents, 86.7%, place the designer/evaluator into an active role during developmental testing. This would logically follow the active role of the test subject; the two roles would naturally complement one another. The most frequent response indicated that if a problem arises within the instructional materials, the designer/evaluator asks "why", pinpoints the problem, and gives just enough help to get the test subject back on track. Raw frequencies indicate that 87.3% of the responses to the questionnaire do not restrict the amount of interaction that can be initiated with the evaluator, in terms of asking questions and making comments to the evaluator.

Recommendation: Most authors place the student into a passive role as far as revision is concerned although a very few authors do suggest that test subjects can be asked to actually make the revisions. Most authors suggest that it is the test subject's behavior that will lead an evaluator into making revisions. Those who take a strictly quantitative position place the test subject in a more restrictive role - subjects learn the material as best they can and then take the test.

Finding: Based on raw frequencies, practitioners agreed with the most frequent recommendations. Test subjects are not actually revising the materials; they are commenting to the designer/evaluator by indicating their problems. This suggests that test subjects are providing a data base for the evaluator on which to ground revisions.

Recommendation: Most authors suggest revising "appropriately" based on the data that has been collected and analyzed. Beyond this, very few authors provide heuristics or algorithms for revision, i.e., certain kinds of errors or problems dictate certain kinds of revisions.

Finding: A major factor affecting the extent of data analysis, for revision purposes, is the resources available (e.g., budget and time constraints). It appears that designers debrief test subjects after their completion of the material, and then tally test results, attitudinal data, and test subject generated comments. The test subject and the designer cooperatively interpret the data for the purposes of identifying problems within the material. Designers then test out the revisions generated from a previous session on the next test subject(s). Strong emphasis is placed on error analysis to find out why individual test subjects make errors. To a lesser extent, designers are comparing pretest and posttest data with a pre-set standard. After having obtained and analyzed their data, 34% of the practitioners indicated that revision was based solely on "obvious", self-evident weaknesses within the instructional materials. Only 11% of the responses indicated that they consistently followed a prescribed algorithm or heuristic for revision. Six percent of the responses indicated that revision was intuitively based. It is interesting to note that the roles that designer and test subject play during the testing session fundamentally affect the revision process. For example, if the test subject is only permitted to learn the material and take the test, revision must be made based on the test data. These findings from the data analysis of the questionnaire clearly identify the need for the research and development of systematic revision procedures to be used when applying an instructional design model to human learning/performance problems.

Recommendation: Authors are vague in prescribing when to terminate developmental testing and revision procedures. Some suggest using performance criteria such as 80/80 but criteria vary between authors and many authors do not precisely specify when to stop testing and revising the instructional product.

Finding: For terminating developmental testing, respondents indicated after some criterion response pattern was being met (pre-set criteria or otherwise). Naturally, project deadline, client, money available, etc., was the over-riding factor. Similarly for the termination of revision, the instructional designers responding to the questionnaire stopped revising when some consistent criteria were being met, e.g., 80/80, 90/90, or after several sessions revealed consistent results.

The following findings were not prompted by recommendations from the literature but are of interest to practitioners. Respondents were asked if both developmental testing and revision procedures were teachable. The response pattern was strikingly similar for the two activities. A large majority of respondents either stated that their procedures could be/are taught to others or they stated that even though their procedures were not documented - they could easily be documented and therefore teachable. This does suggest that more effective teaching can be produced if guidelines for developmental testing and revision can be validated.

Conclusion

Based on the response patterns within the results of the data analysis of the questionnaire, the following represents a "profile" of developmental testing and revision procedures:

1. Select a "representative" sample of test subjects for whom the instruction has been designed.
2. Either conduct several one-to-one sessions and then conduct small group sessions or just conduct small group sessions.
3. Pretest the subjects, administer the learning material, interact with the subjects whenever a problem is noticed in order to locate and define the problem, give just enough help to get the subjects back on track, note attitudinal reactions of the subjects, posttest the subjects, and finally debrief the subjects.
4. Cooperatively analyze and interpret the data with the test subjects.

5. Revise, after the test subject(s) have been debriefed, based on test subjects' behavior during the session.
6. Test out your revisions during the next session.

It should be noted that the above profile does not represent any one particular practitioner or author's suggestions. It represents a unique compilation of practitioner reports based on guidelines from the textbook literature.

Some of the procedures listed above raise some important questions. For the selection of a representative sample of test subjects, it remains unclear as to what characteristics/traits optimize test subjects. There is the issue of test subject ability. Does ability in this context refer to job performance, prior knowlege of the material to be learned, or is it the ability to learn? Clearly more research needs to be conducted on the impact of these and other test subject characteristics that will affect the data. Paralleling this issue, it remains unclear as to how to sample from the target population in order to insure that ideal test subjects are selected. More research is needed on test subject identification procedures.

The second procedure in the profile remains ambiguous. It will be more cost-effective if one-to-one sessions could be eliminated by utilizing only small group sessions. However, one-to-one sessions might be able to contribute unique data unattainable in small group situations. We are presently investigating the differences in the data obtained under one-to-one and small group conditions. A similar issue relates to the generalizability of data, under both one-to-one and small group conditions, to the target population for which the material is being designed. Extremely small samples are used for developmental testing and it remains a statistical issue as to whether valid data can be generated from such small samples. More research is needed in order to hone sampling procedures in order that valid data be obtained.

Another issue can be addressed to the prescribers (textbook authors) of developmental testing and revision procedures. There is very little agreement on their part as to what "revise appropriately" really means because of the tremendous variability of revision procedures. In fact

many of the procedures conflict across authors. From the practitioners' point of view, regardless of what is prescribed, the real-world constraints are such that it appears unreasonable to expect instructional designers to follow revision prescriptions as they presently exist due to the heavy demands they impose on resources. Research-based and developed guidelines that are "manegeable" appear to be needed.

A third phase of the research programme, presently in progress, consists of an experiment designed to investigate and describe the differences between qualitative and quantitative evaluation mechanisms and the implications for obtaining data from test subjects (students) for the purposes of revision. Test subject reliability for providing feedback will also be addressed along with a better description of the differences in the different kinds of data obtained under one-to-one, small group, and both one-to-one and small group conditions for conducting developmental testing sessions.

References

Abedor, A. (1971). Development and validation of a model explicating the formative evaluation process for multi-media self-instructional learning systems. Doctoral dissertation, Michigan State University, Lansing, Michigan, (ERIC ED 064 327, 1972).

Andrews, D. & Goodson, L. (1980). A comparative analysis of models of instructional design. Journal of Instructional Development. 3(4), 2 - 16.

Baker, E. & Schutz, R. (1971). Instructional product development. New York: Van Nostrand Reinhold.

Bell, N. & Abedor, A. (1977). Developing audio-visual instructional modules for vocational and technical training. Englewood Cliffs, N.J.: Educational Technology Publications.

Bjerstedt, A. (1972). Educational technology: Instructional programming and didakometry. New York: McGraw Hill.

Davis, R., Alexander, L., & Yelon, S. (1974). Learning system design: An approach to the improvement of instruction. New York: McGraw-Hill.

Deterline, W. & Lenn, P. (1972). Coordinated instructional systems. Palo Alto, Calif.: Sound Education, Inc.

Dick, W. & Carey, L. (1978). The systematic design of instruction. Glenview, Ill.: Scott, Foresman & Company.

Friesen, P. (1973). Designing instruction. Santa Monica, Calif.: Miller Publishing Company.

Geis, G., Burt, C. & Weston, C. (1984). Instructional development: Developmental testing. A Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, (ERIC ED 243 793).

Geis, G. & Weston, C. (1983). Performance system design. Unpublished manuscript, McGill University, Montreal.

Gropper, G. (1975). Diagnosis and revision in the development of instructional materials. Englewood Cliffs, New Jersey: Educational Technology Publications.

Horn, R. (1966). Developmental testing. Ann Arbor, Michigan: Center for Programmed Learning for Business.

Johnson, R. & Johnson, S. (1975). Toward individualized learning: A developer's guide to self-instruction. Reading, Massachusetts: Addison-Wesley Publishing Company.

Lawson, T. (1974). Formative instructional product evaluation: Instruments and strategies. Englewood Cliffs, New Jersey: Educational Technology Publications.

BEST COPY AVAILABLE

Markle, S. (1978). Designs for instructional designers. Champaign, Ill.: Stipes Publishing Company.

Merrill, D. & Tennyson, R. (1977). Teaching concepts: An instructional design guide. Englewood Cliffs, N.J.: Educational Technology Publications.

SPSSX User's Guide (1983). Chicago, Ill.: McGraw-Hill Book Company.

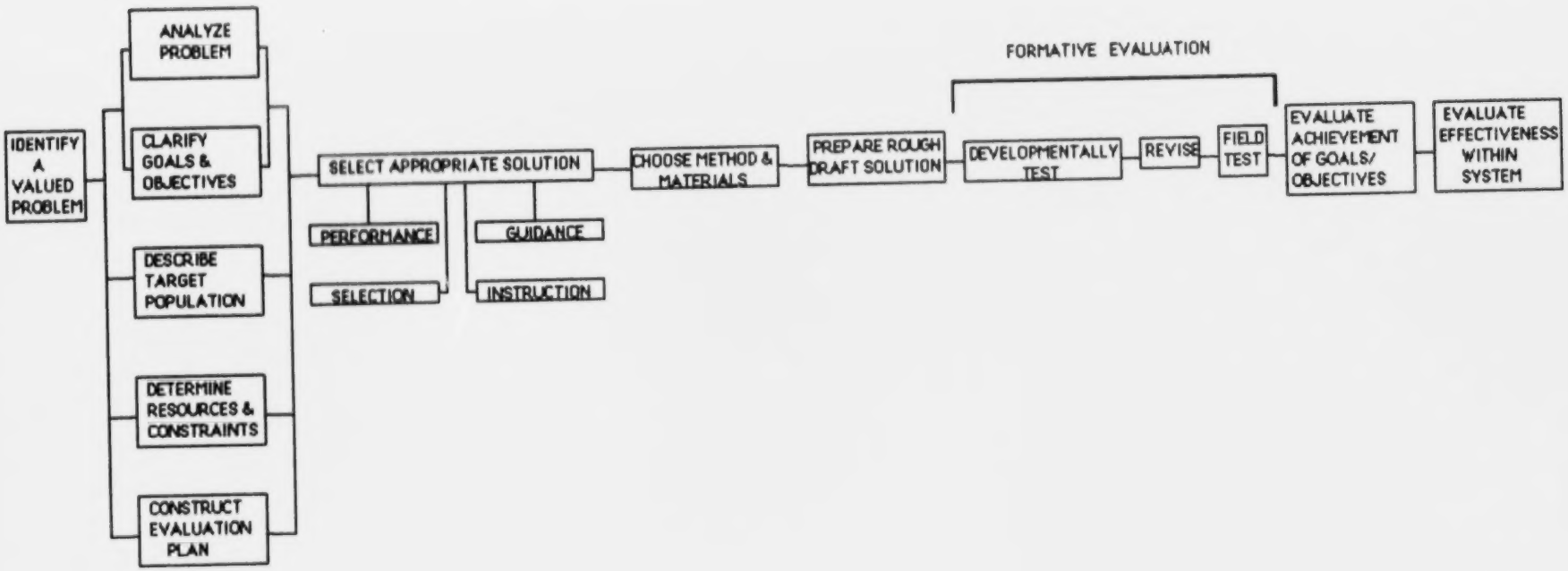
Thiagarajan, S., Semmel, D., & Semmel, M. (1974). Instructional development for training teachers of exceptional children: A sourcebook. Bloomington, Indianna: Center for Innovation in Teaching Handicapped.

Weston, C., Burt, C., & Geis, G. (1984). Instructional development: Revision procedures. A Paper presented at the annual meeting of the American Educational Research Association, New Orleans.

Wydra, F. (1980). Learner controlled instruction. Englewood Cliffs, New Jersey: Educational Technology Publications.

BEST COPY AVAILABLE

Figure 1 A General Model For Instructional Planning (Adapted from Geis & Weston 1983)



BEST COPY AVAILABLE

Table I

Categorization of Guidelines Proposed by Selected Authors for Conducting Developmental Testing Session.

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Abedor (1971)	80%/80% on post test; 80% on reactionnaire	9-12 in small group debriefing	varying abilities (SAT scores)	semi-passive during testing; active during debriefing	passive during testing; active during debriefing	post tests, reactionnaire, tape recordings, observation, debriefing comments	post test, reactionnaire, debriefing	cites Horn
Baker & Schutz (1977)	test until performance variability eliminated	not too small; not too large	-	passive learner	summarizes testing data (developer and tester may be different)	depends on type/format of test	criterion test and "Student Response"	-
Bell & Abedor (1977)	70-80% of students correct on test item; 70% satisfaction on reactionnaire	4-6 in small group debriefing	varying abilities	(see ABEDOR)	(see ABEDOR)	(see ABEDOR)	student revisions (consider cost-benefit); (also: see ABEDOR	-
B.Herstedt (1972)	-	one or more	typical of population; not too timid.	active during testing: thinking aloud individual debriefings; completes pre and post test and reactionnaire	tutors student during testing	pre/post test; reactionnaire; student comments; debriefing.	error analysis; student comments	-
Davis, Alexander & Yelon (1974)	-	-	representation group	-	designs tests, debriefs	achievement tests; reactionnaire	-	-

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Deterline & Lenn (1972)	tutor student until meets objective	3 students; one-to-one	representative of target population	semi-passive during testing; active during debriefing	tutors during testing; debriefs; analyzes results	tape-recorded tutoring (if required); notes; probing;post test	revision algorithm; types of errors	-
Dick & Carey (1978)	-	2 or more for 1 to 1; small group of 10-20	high, average, low in general ability	active, discusses difficulties	active; revises on the spot	notes, revisions made during testing	errors on test items; student comments, revisions, reactions, developer notes.	-
Friesen (1973)	7% error rate on "frame"	one-to-one until success; small group (6); large group (30)	representative of target population	complete materials and tests; responds to developer's probes.	semi-passive during testing; intervenes if required.	notes on testing; analysis of test results.	student testing behaviours and test results.	-
Gropper (1975)	see BASIS OF REVISION						provides numerous decision tables & algorithms implying various rules and strategies.	-
Horn (1966)	-	varies from 1-12, usually 3	representative of target population; include high, average, low.	completes material, writes or expresses comments; incl. confidence level.	introduce testing; give tests; times, revises on spot; takes notes; interviews.	(see ROLE OF DEVELOPER)	student comments; test data; on-the-spot revision.	informal; free from distraction.

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Johnson & Johnson (1975)	70%/90%	2 or 3	anyone	work through material, recording difficulties	revises on spot, if possible.	student revision comment sheet	apparent difficulties	-
Lawson (1974)	-	one-on-one; small group	use below average student	completes materials, tests and questionnaires; criticizes	records all diagnostic information; discusses difficulties	numerous forms for internal and external evaluation	expert input, student opinions and comments, test results.	-
Markle (1978)	acceptable test data	-	testing reveals critical prerequisites	complete materials and tests	analyzes test data (emphasis on valid criterion test)	test data	test results	-
Merrill/ Tennyson (1977)	test results; confidence rating by student	one-on-one, (repeat as often as necessary); then small group (N - 15)	-	go through materials; makes comments; take tests; note confidence one-on-one.	observe student and take notes, clarify points of difficulty	notes by student and developer; on-the-spot revision; pre/ post tests; student confidence ratings.	early revisions based on clinical testing; later revision based on group test data, algorithms for revision.	-

BEST COPY AVAILABLE

	Criterion Performance	Number of Students	Population Characteristics	Role of Student	Role of Developer	Method Data Collection	Basis of Revision	Testing Environment
Thiagarajan/ Semmel/ Semmel (1974)	until consistent results are obtained	"qualitative testing" 4 or 5 in 1-to-1 or group; "quantitative testing - small group.	enthusiastic and verbal students	one-on-one: go through materials, think aloud, request assistance; small group: act as learner	one-on-one-monitors, gives assistance when needed; administer and monitor group session.	student verbal/ non-verbal reactions; interviews; pre/post test group data.	on-the-spot changes; test data for group.	-
Wydra (1980)	-	-	-	learner; makes comments	passive during testing; "inconspicuous", administers session.	tape record session; student written comments; developer comments (written on material).	notes by student and developer; test data	-

BEST COPY AVAILABLE

TABLE 2

RESPONSE RATE FOR THE QUESTIONNAIRE

	TOTAL SAMPLE	%	RESPONDENTS	%
Respondents not involved in Instructional Design	N=32	16%	N=32	25%
Respondents using another approach	N= 5	2%	N= 5	4%
Respondents used in the data analysis	N=90	44%	N=90	71%
Non-respondents	N=79	38%	-	-
Total	**N=206**	**100%**	**N=127**	**100%**

Total Response Rate= 127/206 = 62%

BEST COPY AVAILABLE

TABLE 3

DEMOGRAPHIC BREAKDOWN OF SAMPLES' PRINCIPAL TYPE OF EMPLOYMENT

	Total	Respondents	Non-respondents	Not involved in I.D.
	N=206	N=90	N=79	N=32
	%	%	%	%
Academic/Educational Institution	29%	29%	27%	25%
Consulting/Contracting Firm	20%	21%	22%	25%
Independent Consultant	4%	6%	3%	6%
Private Business/ Industry	19%	20%	4%	9%
Government Agency (Federal,State, City) Other Than Military	7%	8%	18%	9%
Military (As Civilian or Service Personnel)	7%	4%	15%	3%
Utility Company (Such as Power, Communications, etc.)	7%	9%	11%	6%
Medical Facility	4%	2%	-	6%
Other	2%	1%	-	9%
Total	**99%**	**100%**	**100%**	**98%**

Note that some columns do not total 100% due to rounding.

BEST COPY AVAILABLE

TABLE 4

DEMOGRAPHIC BREAKDOWN OF SAMPLES' MOST SIGNIFICANT EMPLOYMENT RESPONSIBILITY

	Total N=206 %	Respondents N=90 %	Non-respondents N=79 %	Not Involved in I.D. N=32 %
Training Director/ Manager	26%	29%	27%	16%
Developer of Training and/or Non-training Solutions	21%	23%	22%	16%
Trainer (Delivery/ Training)	1%	-	3%	-
Developer and Trainer (Equal Responsibilities)	7%	9%	5%	6%
Researcher, Consultant	22%	22%	18%	31%
Teacher/Professor	14%	12%	15%	16%
Other	8%	4%	11%	16%
Total	**99%**	**99%**	**101%**	**101%**

Note that columns do not total 100% due to rounding.

BEST COPY AVAILABLE

APPENDIX 1: QUESTIONNAIRE AND ACCOMPANYING COVER LETTER

Questionnaire on Developmental

Testing and Revision

Please return this questionnaire in the self-addressed envelope enclosed.

c/o Dr. George Geis

Dept. of Educational Psyc.,

Faculty of Education

3700 McTavish St.

McGill University

Montreal, Quebec, Canada

H3A 1Y2

1. In developmentally testing proto-type instructional materials, which of the following general strategies do you usually employ? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ testing with subjects on a one-to-one basis.

___ ☐ testing with subjects on a small group basis.

___ ☐ in addition to testing subjects, I obtain information from subject matter experts or peer reviews.

___ ☐ instead of testing subjects, I obtain information from subject matter or peer reviews.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

BEST COPY AVAILABLE

2. In revising prototype instructional materials, which of the following general strategies do you usually employ? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ revision is based on test data.

___ ☐ revision is based on an attitudinal survey.

___ ☐ revision is based on test subjects' comments.

___ ☐ revision takes place during the testing session.

___ ☐ revision takes place after the testing session.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

3. What aptitudes or levels of performance do you require of test subjects utilized during the developmental testing and revision phase of the instructional design process? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ high level/aptitude performers only.
- ☐ high and low level/aptitude performers.
- ☐ high, medium, and low level/aptitude performers.
- ☐ enthusiastic and highly verbal test subjects.
- ☐ OTHER Please outline briefly)

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

4. What kind of instrument do you usually use to differentiate test subject level/aptitude before the testing session? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ I.Q. Test
- ☐ a standardized aptitude test.
- ☐ past achievement data (e.g., G.P.A., course grades, etc.).
- ☐ job performance measures.
- ☐ previous instructors' opinions.
- ☐ pre-test on material to be taught.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

5. How many sessions and test subjects do you typically involve during developmental testing and revision? PLEASE FILL IN THE BLANK(S).

a) number of one-on-one sessions._____

b) number of small group sessions._____

c) number of test subjects typically used in each small group session._____

OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

6. What do you consider the role of the test subject to be during developmental testing? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ test subject provides pre- and post-test criterion data.

___ ☐ test subject provides attitudinal data.

___ ☐ test subject comments on problems with the instructional material.

___ ☐ test subject comments on problems with the tests.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

7. What do you consider the role of the test subject to be during the revision of the instructional materials? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ the role of the test subjects is to learn as evidenced in pre- and post-test data.

___ ☐ test subjects comment verbally, indicating problems.

___ ☐ test subjects are asked to aggressively determine the causes of their problems.

___ ☐ test subjects are asked to determine the causes of their problems only when the instructional developer notices a problem.

___ ☐ test subjects comment by writing on the instructional material itself or on a revision sheet.

___ ☐ test subjects actually revise instructional material as they work through it.

___ ☐ test subjects actually revise instructional material after they work through it.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

8. Do you see different test subjects as contributing qualitatively different information during successive developmental testing and revision sessions or do you find that test subjects' comments generally replicate each other.

- ☐ different subjects often provide quite different kinds of information.
- ☐ subjects' generally replicate each others comments.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

9. During a developmental testing session, test subjects are: PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ not permitted to ask any questions.

___ ☐ not permitted to make comments.

___ ☐ permitted to but not encouraged to make comments or ask questions.

___ ☐ actually encouraged to ask questions or make comments.

___ ☐ actually taught by the developer when they run into difficulty.

___ ☐ thinking aloud and requesting assistance when required.

___ ☐ recording problems on a revision sheet and/or on the actual instructional materials as they work through the material.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

BEST COPY AVAILABLE

10. What do you consider the role of the developer to be during developmental testing? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ the developer waits, watches, makes notes, and keeps quiet while the test subject works.

___ ☐ if a problem arises, the developer asks "why", pinpoints the problem, and gives just enough help to get the test subject back on track.

___ ☐ the developer personally answers test subject's questions in a tutorial fashion.

___ ☐ the developer notices gaps, predicts confusions, senses response failures, and re-writes items on the spot.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BEST COPY AVAILABLE

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

11. What do you consider the role of the developer to be during the revision of proto-type instructional materials? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

___ ☐ the developer revises on the spot after the test subject has commented.

___ ☐ the developer debriefs a small group of test subjects based on criterion test data, attitudinal survey, and elicited test subject comments. Revision is then based on these debriefing results.

___ ☐ the developer revises based on data generated from teaching the instructional material to the test subject in a tutorial fashion when the subject encounters difficulties.

___ ☐ the developer revises based on a comment sheet that the test subject has used to point out problem areas.

___ ☐ the developer revises on the basis of criterion test data and attitudinal survey after each one-to-one session.

___ ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

55 If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

12. When do you stop developmental testing?
PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ after a pre-determined number of one-to-one sessions.
- ☐ after having debriefed students from one-to-one sessions.
- ☐ after having debriefed students in a small group setting.
- ☐ after criterion test results begin indicating a "success" factor.
- ☐ when a previously set criterion level of achievement and attitudinal response is being met during one-to-one testing.
- ☐ when a previously set criterion level of achievement and attitudinal response is being met from small group debriefing sessions.
- ☐ depends primarily on project deadline dates, client, money available, etc.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

13. When do you stop revising? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(s).

- ☐ I usually set no specified criteria for test subject performance.
- ☐ when less than 7% of the test subjects are having problems with the material.
- ☐ after 4 or 5 sessions until consistent results are obtained.
- ☐ when there are not more than 20% unsatisfactory responses on the attitudinal survey.
- ☐ when 70% of the test subject group achieve 70% or higher on the post-test.
- ☐ wnen 80% of the test subject group achieve 80% or higher on the post-test.
- ☐ OTHER (please outline briefly) _____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

If any answers apply solely or primarily to one stage of developmental testing, indicate so by using this code (which will appear often in the questionnaire). Place the appropriate letter on the line next to your check mark(s).

a. early stage
b. middle stage
c. late stage (just before field testing)

14. How do you obtain your data during <u>developmental testing</u>? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S). AFTER HAVING CHECKED OFF YOUR RESPONSE, PLEASE INDICATE BY PLACING A STAR, ★, NEXT TO THE ITEM(S) PROVIDING THE MOST USEFUL DATA FOR REVISION.

__ ☐ from criterion pre- and post-tests.

__ ☐ from an attitudinal survey.

__ ☐ from the probing (during the test session) of test subjects.

__ ☐ by recording (by note-taking, audio tape, or video tape) information.

__ ☐ from inferences based on observed test subject behavior.

__ ☐ by debriefing test subjects (after the test session).

__ ☐ from test subject notations written on the instructional materials.

__ ☐ from test subject comments placed on a revision comment sheet.

__ ☐ from the notes that were taken on examples, illustrations, information added, and changes to sequences made while the test subject was tutored through the instructional material.

__ ☐ from spontaneous verbal comments that the test subjects have made during the test session.

__ ☐ OTHER (please outline briefly) _____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

15. How do you analyze the data gathered from developmental testing sessions? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ the extent of the analysis depends on the resources available (e.g., budget and time constraints).
- ☐ by comparing the cumulative effects of revisions based on a number of test sessions (e.g.,by testing out the revisions generated from a previous session on the next test subject(s).
- ☐ by debriefing test subjects after their completion of the material, thus providing tallied test results, attitudinal data, and test subject generated comments. The test subjects and the developer cooperatively analyze the data.
- ☐ by comparing pre- and post-test data with a pre-set standard.
- ☐ with strong emphasis placed on error analysis to find out why individual test subjects make errors.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

16. For purposes of revision after you have obtained information from developmental testing, do you: PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

- ☐ follow a prescribd algorithm or heuristic, e.g., certain kinds of errors dictate certain changes.
- ☐ revise on the basis of "obvious", self-evident weaknesses within the instructional material.
- ☐ hard to say; it is intuitive.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

17. We recognize that many instructional developers have evolved sophisticated procedures over the years, for developmental testing and revision. We are interested in learning how explicit and teachable your procedures are, i.e., can they be taught to others? PLEASE CHECK, ☑, THE MOST APPROPRIATE RESPONSE(S).

A. With regard to developmental testing?

- ☐ my procedures are well documented in the literature or in handouts, and could be/are taught to others.
- ☐ my procedures could easily be documented and taught to others but have not been.
- ☐ it would be difficult for my procedures to be documented or taught to others.
- ☐ my procedures can not be (or are not likely to be) documented or generalized because of varying local conditions. Each project, test subject, context, etc., is different and requires a different strategy.
- ☐ OTHER (please outline briefly)

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

BEST COPY AVAILABLE

B. With regard to revision?

- ☐ my procedures are well documented in the literature or in handouts, and could be/are taught to others.
- ☐ my procedures could easily be documented and taught to others but have not been.
- ☐ it would be difficult for my procedures to be documented or taught to others.
- ☐ my procedures can not be (or are not likely to be) documented or generalized because of varying local conditions. Each project, test subject, context, etc., is different and requires a different strategy.
- ☐ OTHER (please outline briefly)

_____

_____

_____

_____

_____

BASED ON EXPERIENCE, RESEARCH, OR ADVICE, WHAT HAS PROMPTED YOU TO UTILIZE YOUR SELECTION(S)?

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

We would very much appreciate your enclosing or providing a reference to any documentation on developmental testing or revision that you have generated and/or use.

ENCLOSING ○

NOT ENCLOSING ○

18. You have described your methods of developmental testing and revision in the previous items. We recognize that everyday constraints may impose upon you procedures that you consider less than ideal. Would you take a couple of minutes to describe below what you consider to be the ideal procedures for developmental testing and revision, considering, if possible, the variables noted in the questionnaire (e.g., number of test subjects, role of the developer, etc.?)

BEST COPY AVAILABLE

THERE IS NO NEED TO COMPLETE THIS ITEM IF YOU HAVE FILLED OUT THE REST OF THE QUESTIONNAIRE

19. As indicated in the cover letter, you are filling out this page because you have a different method for developing and improving instructional material than the one we described (i.e., developmental testing and revision). Would you kindly take a minute or two to describe briefly how you do it?

20. Please just check the box if you wish to be kept up to date on the research project.

I would like to receive:

- ☐ AERA papers
- ☐ Results of this study
- ☐ Both

Department of Educational Psyc.,
Faculty of Education, McGill University,
3700 McTavish St., Montreal, P.Q.

July, 5, 1984.

Dear Colleague,

We ask your assistance in carrying out a research project on systematic instructional design. The focus of the study is on developmental testing and revision. We have compiled, through a search of the literature, a series of suggestions and prescriptions for carrying out developmental testing and revision. The enclosed questionnaire is designed to obtain from experts like you, opinions on these suggested procedures.

The goal of this phase of the research is to determine the amount of agreement on and the rationales for using the procedures outlined in the questionnaire. We will follow up this phase with several experimental studies.

We will be happy to send you a copy of the results of this project (if you care to do some background reading on the research, we will send you copies of two 1984 AERA presentations; just check the box at the end of the questionnaire). We hope to present the research at NSPI meetings and in Performance and Instruction.

By "developmental testing" we mean a procedure of trying out instructional units (often self-instructional) in draft form with a small number of subjects (learners or performers) in order to obtain information to be used in revising and improving the instruction. It is done "in-house" before the learning materials are in final form and ready to be released to the general learner population.

In contrast, field testing generally occurs later in the development process; it is often conducted at the site at which the materials will finally be used; it usually involves a large population of learners; it assumes that major revisions have already been carried out based on developmental testing and that field testing is a final validating procedure.

By "revision" we mean a diagnostic and prescriptive process, based on developmental testing, in order to produce instructional materials that meet set performance criteria.

We realize that if you are involved in developmental testing and revision procedures which more or less fit these descriptions, in the real world your practices are likely to vary under different circumstances. While responding to the questionnaire, would you indicate what you do most frequently? We also realize that your most frequent practices may not be "ideal" or the most effective. We have provided (on the next to last page of the questionnaire) an opportunity to describe the way you would test and revise materials if you were not operating under practical constraints. We are asking in the questionnaire for generalizations which would help in building general "rules" or "heuristics". Consequently we would appreciate your doing the best you can to generalize even if you feel uncomfortable doing so.

If you do not engage in developmental testing as we have defined it, please turn to the last page of the questionnaire and complete it.

If you are not involved in designing systematic instruction, please return the questionnaire so that we can maintain complete records.

Some of the items in the questionnaire may appear to overlap and be redundant. This is due to the close relation between developmental testing and revision.

We thank you in advance for your cooperation and look forward to receiving your important contribution to the research project.

Sincerely Yours,

Dr. George Geis

George Geis

Dr. Cynthia Weston

Cynthia Weston

Mr. Charles Burt

Charles Burt